

On a Minimax Problem for Ovals

Rolf Walter

1. Introduction

Let (X, d) be a bounded metric space. A fundamental invariant of X is the *diameter*

$$(1.1) \quad \text{diam}(X) := \sup_{p, q \in X} d(p, q).$$

Besides $\text{diam}(X)$, one can form the quantity

$$(1.2) \quad \delta(X) := \inf_{p \in X} \sup_{q \in X} d(p, q),$$

which, in general, is distinct from $\text{diam}(X)$.

$\delta(X)$ is the smallest radius of balls covering X with centers in X . Despite this obvious meaning, the invariant $\delta(X)$ seems rather untouched in the geometric literature. In the area of approximation theory, $\delta(X)$ occurs as the relative Chebyshev radius of X w.r.t. X itself, i.e. $\delta(X) = r(X, X)$ in the terminology of *Amir/Ziegler* [1980]. The size of $\delta(X)$ might have an impact on the shape of the space X . In particular, for convex hypersurfaces in \mathbb{R}^n , there could exist relations to the many other geometric invariants of such objects.

In this paper we discuss, for plane convex curves X , an isoperimetric type inequality, relating $\delta(X)$ to the perimeter of the curve. Due to the minimax character of $\delta(X)$, its handling resists the usual principles for extremal questions. It will be shown that the smooth case of the inequality can be reduced to the polygonal case via approximation. However, for polygons, there is the additional problem of the high dimensionality of the set of vertices. So, in general, we only can offer conjectures. Definite solutions are possible for restricted classes of curves. Even for short polygons, there already arises a sort of ‘magic kites’ which show that squares are definitely not the extremal figures.

In the general context, it is easy to see that the function of farthest distance

$$(1.3) \quad \mu : X \longrightarrow \mathbb{R}, \quad \mu(p) := \sup_{q \in X} d(p, q)$$

is Lipschitz-continuous with Lipschitz-constant 1. In fact, from the triangle inequalities for points $p_1, p_2, q \in X$:

$$(1.4) \quad d(p_1, q) - d(p_1, p_2) \leq d(p_2, q) \leq d(p_1, q) + d(p_1, p_2)$$

one derives, by taking suprema w.r.t. q

$$(1.5) \quad \mu(p_1) - d(p_1, p_2) \leq \mu(p_2) \leq \mu(p_1) + d(p_1, p_2),$$

hence

$$(1.6) \quad |\mu(p_1) - \mu(p_2)| \leq d(p_1, p_2).$$

If X is compact then the suprema in (1.1) and (1.2) may be replaced equivalently by the maxima, because $d(p, q)$ is continuously dependent on the argument pair. By the continuity of μ , the same is true for the infimum in (1.2). More precisely we can state:

1.1. Lemma. *Let X be a compact metric space X . Then we have*

$$(1.7) \quad \delta(X) = \min_{p \in X} \max_{q \in X} d(p, q).$$

In particular, there are $p_0, q_0 \in X$ with

$$(1.8) \quad \delta(X) = d(p_0, q_0) \geq d(p_0, q) \quad \text{for all } q \in X.$$

Conversely, if $p_0, q_0 \in X$ are such that (1.8) holds, then the function $\mu(p) := \sup_{q \in X} d(p, q) = \max_{q \in X} d(p, q)$ attains its minimum at p_0 , and the function $q \mapsto d(p_0, q)$ attains its maximum at q_0 .

Proof. The function μ assumes its minimum at some point $p_0 \in X$, and then the function $r \mapsto d(p_0, r)$ assumes its maximum at some point $q_0 \in X$. Thus

$$(1.9) \quad \max_{r \in X} d(p, r) \geq \max_{r \in X} d(p_0, r) = d(p_0, q_0) \geq d(p_0, q) \quad \text{for all } q \in X.$$

This implies by switching to the infimum w.r.t. p

$$(1.10) \quad \delta(X) = \min_{p \in X} \max_{r \in X} d(p, r) = d(p_0, q_0) \geq d(p_0, q) \quad \text{for all } q \in X.$$

This proves (1.7) and (1.8) for this specific pair (p_0, q_0) .

Conversely, if p_0, q_0 is *any* pair in X , satisfying (1.8), then $\min_{p \in X} \mu(p) = d(p_0, q_0) \geq d(p_0, q)$ for all $q \in X$. This implies $\mu(p_0) = \max_{q \in X} d(p_0, q) = d(p_0, q_0)$, hence $\min_{p \in X} \mu(p) = \mu(p_0)$. The second relation in (1.8) shows the remainder of the assertion. \square

We call a pair of points $p_0, q_0 \in X$ *distinguished* if it fulfills (1.8). Equivalent to this is that the function μ has the minimum point p_0 and the function $q \mapsto d(p_0, q)$ has the maximum point q_0 . So p_0, q_0 can well be viewed as points where the minimax-invariant $\delta(X)$ attains its value. Observe that the roles of p_0, q_0 are not interchangeable. It is possible that there are more than one pair of distinguished points, e.g. if X has isometric symmetries. In the terminology of approximation theory, the first point p_0 of a distinguished pair p_0, q_0 is the same as a relative Chebyshev center.

Now let X be a simply closed convex curve Γ in \mathbb{R}^2 equipped with the metric d , restricted from the standard Euclidean metric in \mathbb{R}^2 . Assume that the domain Ω enclosed by Γ is non-void. Denote by $C := \Omega \cup \Gamma$ the corresponding convex body and by $L(\Gamma)$ its perimeter.

A distinguished pair of points p_0, q_0 at which δ assumes its value according to (1.8) determines the related chord as the segment from p_0 to q_0 . Thereby, q_0 is the maximal point of the function $q \mapsto d(p_0, q)$. Such a chord (which of course is contained in C) will be called *distinguished*. Thus, each closed convex curve determines certain peculiar chords which, without doubt, have a geometric meaning. However, at this time, only few is known on this. E.g. for a C^2 -curve, it is open how the distinguished chords can be determined. For convex polygons we shall detect some details of distinguished chords in section 5. For instance, the second endpoint of a distinguished chord must be a vertex, and there exists a finite algorithm for the finding of all distinguished chords.

Since each circle of radius $\delta(\Gamma)$ around a relative Chebyshev center p_0 encloses Γ , one has

$$L(\Gamma) \leq 2\pi\delta(\Gamma),$$

by the monotony of the boundary volume for convex sets (see *Bangert* [1981] for a rather general version of the monotony). In the opposite direction we state the

Conjecture 1: *For any closed convex curve Γ in the plane \mathbb{R}^2 , the inequality*

$$(1.11) \quad L(\Gamma) \geq \pi \cdot \delta(\Gamma)$$

holds true.

This conjecture is supported by the results of this paper and by numerous experiments on further classes of curves.

Returning to the definition of $\delta(\Gamma)$, one can state the conjecture (1.11) as follows: *Let δ_0 be a positive constant such that, for any point $p \in \Gamma$, there is a point $q \in \Gamma$ with $d(p, q) \geq \delta_0$. Then $L(\Gamma) \geq \pi \cdot \delta_0$.* If this were true in general it would answer a question recently raised in the field of Reverse Engineering (*Wortmann* [2004]).

2. Curves nearby Circles

By this we will understand that all the curvature centers fall inside the curve:

2.1. Theorem. *Let the closed convex curve Γ in \mathbb{R}^2 be regular and of class C^2 . Assume that all curvature centers of Γ lie in the interior Ω . Then we have*

$$(2.1) \quad L(\Gamma) \geq \pi \cdot \delta(\Gamma).$$

Equality in (2.1) holds if and only if Γ is of constant breadth.

For any $p \in \Gamma$ there is at least one point $q_1 \in \Gamma$ where the function $q \mapsto d(p, q)$ reaches its maximum. The closed segment $[pq_1]$ is orthogonal to the tangent of Γ at q_1 . However, it is not clear whether q_1 is unique or how it depends on p . We therefore consider an eventual inverse of this map, defined as follows: For any $q \in \Gamma$ let $\nu(q)$ be the inward normal of Γ at q (counted as a half-ray starting at q). It cuts Γ in a unique second point $\Phi(q)$, and the open segment $]q, \Phi(q)[$ is contained in the interior Ω . The map

$$(2.2) \quad \Phi : \Gamma \longrightarrow \Gamma$$

will play a decisive role in our arguments.

The orientation of Γ shall always be such that the interior Ω lies to the left. At any point $p \in \Gamma$ we then have a unique inner unit normal vector $N(p)$ of Γ , and for any positively oriented tangent vector w of Γ at p , the pair $(w, N(p))$ will be positively oriented w.r.t \mathbb{R}^2 .

Now, let the segment $[p, q]$ be any *chord* of Γ , i.e. p, q are distinct points on Γ , and $]p, q[$ belongs to the interior Ω . The straight line g spanned by p, q decomposes Γ into two arcs

which lie on different sides of g . For each of these arcs, p and q are the points where the arc enters or leaves (transversally) the half plane of g which contains the arc. So, if u , resp. v are tangent vectors of Γ at p , resp. q then u, v are equally oriented w.r.t. Γ if and only if they point into different sides of g .

2.2. Lemma. *Under the assumptions of Theorem 2.1, the map Φ in (2.2) is an orientation-preserving C^1 -diffeomorphism of Γ onto itself.*

Proof. Let $\gamma : [0, \ell] \rightarrow \mathbb{R}^2$ be a positive parametrization of Γ by arclength s such that γ is injective on $[0, \ell[$ and $\gamma(0) = \gamma(\ell)$. (If necessary, we may extend γ periodically to \mathbb{R} with simple period ℓ .) For the parametrizations of the unit tangent and normal vectors T and N we have the structure equations of Darboux

$$(2.3) \quad \gamma' = T, \quad T' = \kappa N, \quad N' = -\kappa T,$$

κ being the curvature function. The curvature centers are given by

$$(2.4) \quad z = \gamma + \frac{N}{\kappa},$$

where κ is always positive because the curvature centers are in Ω (thus finite): Γ is strictly convex. For the Φ -images of the curve points we have a representation of the form

$$(2.5) \quad \Phi(\gamma(s)) = \gamma(s) + D(s)N(s) = \gamma(\varphi(s)),$$

where $D(s)$ is unique, namely equal to the distance $d(\gamma(s), \Phi(\gamma(s)))$, and $\varphi(s)$ is unique modulo ℓ . In order to show the differentiability of these quantities we consider the C^1 -function

$$(2.6) \quad F(s, \xi, \eta) := \gamma(s) + \xi N(s) - \gamma(\eta)$$

and apply the implicit function theorem to the equation $F(s, \xi, \eta) = 0$. The partials of F w.r.t. ξ and η are $N(s)$ and $-\gamma'(\eta)$ which for $\eta = \varphi(s)$ are linearly independent by the transversality mentioned above. Thus D is of class C^1 , and the same is true for Φ by (2.5). Then φ can be chosen continuously, satisfying $\Phi \circ \gamma = \gamma \circ \varphi$, so in fact being of class C^1 . (Here it is necessary to view all functions on the whole of \mathbb{R} .)

The assumption on the curvature centers is expressed by

$$(2.7) \quad D(s) > \frac{1}{\kappa(s)} \quad \text{for all } s.$$

Differentiating the right equation in (2.5) and inserting according to (2.3) we obtain

$$(2.8) \quad (1 - D\kappa)T + D'N = (\gamma' \circ \varphi) \cdot \varphi'$$

and by scalar multiplication with γ'

$$(2.9) \quad (1 - D\kappa) \langle \gamma', \gamma' \rangle = \langle \gamma' \circ \varphi, \gamma' \rangle \cdot \varphi'.$$

For each s , $\gamma'(s)$ is a normal vector of the chord $[\gamma(s), \Phi(\gamma(s))]$, so the two scalar products in (2.9) have different signs. On the other hand, $1 - D\kappa$ is always negative by (2.7); thus, by (2.9), φ' is always positive, hence φ orientation-preserving.

This proves that Φ is an immersion of the compact manifold Γ into Γ , hence a covering projection (*Kobayashi/Nomizu* [1963], Cor. 4.7), say with $m \geq 1$ leaves. Γ is homeomorphic to the sphere \mathbb{S}^1 . Since Φ is orientation-preserving, its mapping degree also equals m . If we had $m > 1$ then, by the Hopf/Lefschetz fixpoint theorem, there would exist a fixpoint of Φ what is not possible. So Φ is injective and has all properties, as asserted. \square

Proof of Theorem 2.1

We continue in the setting of the last proof. The mapping $\Phi \circ \gamma = \gamma \circ \varphi$ is now called $\tilde{\gamma}$ and viewed on $[0, \ell]$. As γ itself, $\tilde{\gamma}$ is a positive parametrization of Γ , injective on $[0, \ell[$ with $\tilde{\gamma}(0) = \tilde{\gamma}(\ell)$. By (2.5), $\tilde{\gamma}$ and its derivative are given by

$$(2.10) \quad \tilde{\gamma} = \gamma + DN, \quad \tilde{\gamma}' = (1 - D\kappa)T + D'N.$$

This implies for the length of Γ :

$$(2.11) \quad \ell = \int_0^\ell |\tilde{\gamma}'| = \int_0^\ell \sqrt{(D\kappa - 1)^2 + D'^2} \geq \int_0^\ell (D\kappa - 1) = \int_0^\ell D\kappa - \ell.$$

For each $\tilde{\gamma}(s) =: p$, the $\gamma(s)$ is the only point $q \in \Gamma$ where the distance $d(p, q)$ is maximal. Namely, if $q_1 \in \Gamma$ were a point $\neq \gamma(s)$ with maximal distance from p then the segment $[p, q_1]$ were orthogonal to Γ at q_1 , hence $\Phi(q_1) = p$, in contradiction to the injectivity of Φ . This implies

$$(2.12) \quad D(s) = \max_{q \in \Gamma} d(\tilde{\gamma}(s), q) \geq \delta(\Gamma).$$

Thus from (2.11) we infer, using finally the theorem on turning tangents:

$$(2.13) \quad 2\ell \geq \delta(\Gamma) \int_0^\ell \kappa = \delta(\Gamma) \cdot 2\pi.$$

This is the assertion (2.1).

If we have equality in (2.1) then we deduce from (2.11) to (2.13) that $D = \delta(\Gamma)$. For constant D , (2.10) implies $\tilde{\gamma}' = (1 - D\kappa)T$, so the tangents of Γ at $\gamma(s)$ and $\tilde{\gamma}(s)$ are parallel and $d(\gamma(s), \tilde{\gamma}(s)) = D = \text{const.}$ This shows that Γ has constant breadth $\delta(\Gamma)$.

On the other hand, if a closed regular and convex C^2 -curve Γ has constant breadth a then, by well-known properties of such ovals, the curvature centers are situated in the interior and $\delta(\Gamma) = a = L(\Gamma)/\pi$; see e.g. *Bonnesen/Fenchel* [1934], Sect. 63, p. 128. \square

3. Approximating $\delta(\Gamma)$

An important question is whether the minimax-invariant $\delta(\Gamma)$ could be calculated via polygonal approximations. Theorem 3.2 below provides a result in this direction. It contains an error estimate which may be useful for numerical purposes. But above all, eventual improvements of the inequality (2.1) are thus reduced to the polygonal case.

First, we observe two simple facts:

3.1. Lemma.

Let X, Y be bounded subsets of a metric space Z (with the restricted metrics).

- (i) Let ε be a positive constant. If there exists a surjective map $\varphi : X \rightarrow Y$ such that $d(p, \varphi(p)) \leq \varepsilon$ for all $p \in X$ then $\delta(X) \leq \delta(Y) + 2\varepsilon$.
- (ii) If there is a surjective map $\varphi : X \rightarrow Y$ satisfying $d(p, q) \leq d(\varphi(p), \varphi(q))$ for all $p, q \in X$, then $\delta(X) \leq \delta(Y)$.

Proof.

For (i): For any points $p, q \in X$ we have

$$(3.1) \quad d(p, q) \leq d(p, \varphi(p)) + d(\varphi(p), \varphi(q)) + d(q, \varphi(q)) \leq d(\varphi(p), \varphi(q)) + 2\varepsilon,$$

thus, taking suprema w.r.t. q for fixed p ,

$$(3.2) \quad \sup_{q \in X} d(p, q) \leq \sup_{q \in X} d(\varphi(p), \varphi(q)) + 2\varepsilon = \sup_{q_1 \in Y} d(\varphi(p), q_1) + 2\varepsilon,$$

since φ is surjective. For the same reason, taking infima w.r.t. p , this yields the assertion.

For (ii): The assertion is immediately clear from the hypothesis, taking first suprema w.r.t. q and then infima w.r.t. p , observing that $\varphi(q)$ and $\varphi(p)$ run through all of Y . \square

In case of an oriented closed convex and regular C^2 -curve Γ , consider a chord $[p, q]$ and orient the line g connecting p, q in such a way that its right half plane H_g^+ is entered by the curve at p and left at q . The convex body $C := \Omega \cup \gamma$ then cuts H_g^+ in a bi-angle with inner angles $\alpha \in]0, \pi[$ at p and $\beta \in]0, \pi[$ at q . Let Γ_g^+ denote the oriented part of Γ in H_g^+ . The Gauss/Bonnet theorem for this bi-angle says

$$(3.3) \quad \alpha + \beta = \int_{\Gamma_g^+} \kappa \, ds,$$

hence

$$(3.4) \quad \alpha + \beta \leq \max_{\Gamma_g^+} \kappa \cdot L(\Gamma_g^+).$$

In particular, the bi-angle is acute (i.e. α, β both less than $\pi/2$) if $L(\Gamma_g^+)$ is small enough.

Observe that, for an acute bi-angle, the tangents of Γ at p, q intersect in a point r of H_g^+ and that Γ_g^+ is contained in the triangle with vertices p, q, r . Furthermore, the arc Γ_g^+ lies

‘schlicht’ over the side $[p, q]$, i.e. the straight lines orthogonally cutting the points z of $[p, q]$ hit Γ_g^+ exactly once, say at $h(z)$. The distance $d(z, h(z))$ is less than the height h_0 of that triangle over $[p, q]$. By elementary trigonometry, $h_0 = d(p, q)/(\cot \alpha + \cot \beta)$.

3.2. Theorem (on approximation).

Let Γ be an oriented closed and strictly convex C^2 -curve in \mathbb{R}^2 , and let k be an upper bound of its curvature function $\kappa : \Gamma \rightarrow \mathbb{R}$. Let P be a polygon inscribed in Γ and λ an upper bound of the arclength on Γ between all consecutive vertices of P such that

$$(3.5) \quad \lambda < \frac{\pi}{2k}.$$

Then we have for the minimax-invariants of Γ and P

$$(3.6) \quad \delta(P) \leq \delta(\Gamma) \leq \delta(P) + \lambda \cdot \tan(k\lambda).$$

Thus, if λ tends to 0, the minimax-invariant $\delta(P)$ tends to $\delta(\Gamma)$ with an error term of order $\mathcal{O}(\lambda^2)$.

Proof. The vertices of the polygon P on Γ shall follow each other according to the orientation of Γ (P itself is convex with interior in Ω). By (3.4), (3.5), all bi-angles cut out by the edges of P on C are acute having inner angles $\leq k\lambda$. Let $\psi : \Gamma \rightarrow P$ be the metric projection onto P , restricted to Γ , and $\varphi : P \rightarrow \Gamma$ the map whose restriction to any edge $[p, q]$ of P is the map h , defined before. In fact, φ and ψ are inverse to each other. It is well known that ψ is Lipschitz-continuous with Lipschitz-constant 1 (e.g. *Phelps* [1957], Thm. 5.1), hence $d(p, q) \leq d(\varphi(p), \varphi(q))$ for all $p, q \in \Gamma$. On the other hand $d(p, \varphi(p)) \leq \lambda/(2 \cot k\lambda)$. Thus, Lemma 3.1 applies and yields (3.6). \square

4. The Invariant δ for Polygons

By the approximation theorem 3.2 it suffices to prove the desired inequality (2.1) for convex polygons. In this case, the invariant δ can be calculated in finitely many steps. The following observations will show this. Denote by $P = \Gamma$ a convex polygon in \mathbb{R}^2 , by $\Omega \neq \emptyset$ its inner region, and by $C := \Omega \cup \Gamma$ its closure. For $u, v \in \mathbb{R}^2$ we denote by $[u, v]$ the closed segment from u to v , inverting the brackets at an endpoint which does not belong to the segment, e.g. $[u, v] := [u, v] \setminus \{v\}$.

4.1. Lemma. *Given a point $x \in \mathbb{R}^2$, the following holds true: If a point $y \in P$ has maximal distance from x , then y is necessarily a vertex of P .*

Proof. Of course, $y \neq x$. If y would lie in the interior of an edge E then all points of P would lie in the closed disk $\overline{\mathbb{B}}(x, r)$ with radius $r := |x - y|$. On the other hand, the edge E is tangent to the circle $S(x, r)$, by the maximality condition. This contains a contradiction, because a circle tangent minus the point of contact is always outside the circle. \square

Henceforth, a point $y \in P$ in maximal distance to an $x \in \mathbb{R}^2$ will be called a *farthest* point from x . For given $x \in \mathbb{R}^2$ there may exist more than one farthest vertex from P . This happens iff x lies on a perpendicular bisector of two vertices. The set of all perpendicular bisectors of any two vertices of P splits \mathbb{R}^2 in certain domains. More precisely: The complement of all these perpendicular bisectors is an open and dense subset of \mathbb{R}^2 with finitely many connected components. These components arise as the possible intersections of open half spaces of all perpendicular bisectors (as far those intersections are nonvoid). For N perpendicular bisectors we thus have at most 2^N such connected components.

If $x \in \mathbb{R}^2$ owes exactly one vertex $y \in P$ as farthest point from x then all vertices $\neq y$ have a smaller distance to x . By the continuity of the distance function, this is retained when x varies in a small neighborhood. So the set of the $x \in \mathbb{R}^2$ which have y as a unique point of maximal distance on P is open. This implies that all points in one of the connected components from above have the same vertex as a point of maximal distance.

If x varies only in P then this holds accordingly: The intersection points of all perpendicular bisectors with P form a partition of the polygon P , while P is homeomorphic to the unit circle \mathbb{S}^1 . This means: If M_1, \dots, M_K are these intersection points on P , sorted in the positive mathematical sense along P , then the complement $P \setminus \{M_1, \dots, M_K\}$ is composed of the open and connected sets $S_1 := M_1M_2, \dots, S_{K-1} := M_{K-1}M_K$, $S_K := M_KM_1$, and the points of each such set all have the same uniquely determined farthest vertex. Here, for $M \neq M'$ in P , denotes MM' the part of P which traces from M to M' in the chosen orientation of P (without M and M'). In the sequel, the points M_1, \dots, M_K will be called the *section points*, and the sets S_1, \dots, S_K will be called the *sections* of P . For a section S we denote by $y(S)$ the unique farthest vertex, common to all $x \in S$.

It may very well happen that a section S contains vertices of P . So, a section is not necessarily part of an edge. If one refines P by adding the vertices, then there arise new sections which then lie on edges of P . These new sections shall be called *refined sections*. The points of a refined section S' then still have the same unique farthest vertex, namely the one, which belongs to the section S in which the subsection S' is contained. We then write $y(S') := y(S)$.

The above function μ may be determined on any section S : One has $\mu(x) = d(x, y(S))$ for all $x \in S$ since $d(x, q) \leq d(x, y(S))$ for all $q \in P$. By the continuity of μ and d we still have

$$(4.1) \quad \mu(x) = d(x, y(S)) \quad \text{for } x \in \overline{S}.$$

Herein, \overline{S} denotes the closure of S , i.e. S together with its two endpoints. This implies that $y(S)$ is still a farthest point from the two endpoints. However, such an endpoint owes at least one more farthest vertex $\neq y(S)$ because it lies on a perpendicular bisector.

The minimum of the function μ on a (refined) section S equals the minimum of the function $x \mapsto d(x, y(S))$ for $x \in S$ where $y(S)$ does not depend on these x . If, generally, the distance of a point z from a set T is denoted by

$$(4.2) \quad d_T(z) := \min \{d(p, z) \mid p \in T\}$$

then consequently

$$(4.3) \quad \min_S \mu = \min_{\overline{S}} \mu = d_S(y(S)) = d_{\overline{S}}(y(S)).$$

In case of a refined section S the set \overline{S} is a closed segment in \mathbb{R}^2 whose distance from $y(S)$ can be determined elementarily. Namely, if a, b are the endpoints of the open segment S , and if $z \in \mathbb{R}^2$ is arbitrarily given then the nearest point z^* for z on \overline{S} is given by

$$(4.4) \quad z^* = \begin{cases} a + t(b - a) & \text{for } 0 < t < 1 \\ a & \text{for } t \leq 0 \\ b & \text{for } t \geq 1 \end{cases} \quad \text{with} \quad t := \frac{\langle z - a, b - a \rangle}{|b - a|^2}.$$

The cases depend from whether the orthogonal projection of z onto the straight line spanned by S belongs to S (first case) or not (second and third case). In the first case, z^* is the footpoint and the formula follows from the requirement that $z - z^*$ is orthogonal to $b - a$. In the second and third case, z^* is the nearest point for z on S . From this, the distance of z to S is calculated as $d_S(z) = |z - z^*|$, explicitly

$$(4.5) \quad d_S(z) = \begin{cases} \frac{|(z - a) \wedge (b - a)|}{|b - a|} & \text{for } 0 < t < 1 \\ |z - a| & \text{for } t \leq 0 \\ |z - b| & \text{for } t \geq 1. \end{cases}$$

where

$$(4.6) \quad |u \wedge v| := \sqrt{|u|^2 \cdot |v|^2 - \langle u, v \rangle^2}.$$

Remark. In (4.4), the formula for z^* in the first case can also expressed by

$$(4.7) \quad z^* = \frac{\langle z - b, a - b \rangle a - \langle z - a, a - b \rangle b}{|a - b|^2}.$$

The global minimum δ of the function μ is now computable from these finitely many dates:

4.2. Lemma. *If S_1, \dots, S_K are the sections of P then*

$$(4.8) \quad \delta(P) = \min \{d_{\overline{S}_1}(y(S_1)), \dots, d_{\overline{S}_K}(y(S_K))\}$$

The same formula holds if the S_k run through all refined sections of P (then with different K).

Proof. Generally one can say:

Let $X = X_1 \cup \dots \cup X_K$ be a finite covering of a set X and $f : X \rightarrow \mathbb{R}$ be a real function. If there exist the minima of all restrictions $f|_{X_k}$ in \mathbb{R} then there also exists the minimum of f in \mathbb{R} , and one has

$$(4.9) \quad \min f = \min \{\min f|_{X_1}, \dots, \min f|_{X_K}\}.$$

This implies the assertion. □

For a convex polygon, any distinguished chord ends in a vertex (Lemma 4.1).

Eqn. (4.8) allows, in interaction with Eqn. (4.5), the calculation of the minimax invariant δ in a finite number of steps. If one also calculates, in each case, the nearest points $y(S_k)$ on the sections by (4.4) or (4.7) then all distinguished chords can accordingly be obtained in finitely many steps.

5. The Algorithm for Polygons

Here we shall justify and describe in detail the above-mentioned algorithm. Given is only the list of vertices of the polygon P in the order along P , counted mathematically positive. The aim is the determination of the invariant $\delta(P)$ and of the set of all distinguished chords. In the notation, the dependence on P will be suppressed, so instead of $\delta(P)$ we write δ , etc.

The sorting of all points of interest on P follows the *arclength*, measured from the first entry of the list in the counter clockwise manner. In particular, all the arclength parameters of the vertices can be calculated. The arclength parameter of the last entry (identical with the first) is then the perimeter L of P .

Calculation of δ

This takes place with Eqn. (4.8).

First of all, the perpendicular bisectors of *all* pairs of vertices must be cut with P . (The perpendicular bisectors of the edges don't suffice.) Initially, the intersection points don't appear in the correct order along P . However, the arclength parameters of these points can be determined edge-wise. For a better performance one may consider that one of the perpendicular bisectors cuts the polygon in exactly two points. If the perpendicular bisector belongs to an edge then one of these two points the midpoint of the edge. In addition, one may arrange the section points in the order of increasing arclength. With this, the sections can be treated successively, by adding accordingly the arclength parameters of the vertices. In fact, one must work with the refined sections, since the formulae (4.4) to (4.7) refer to a segment but not to an open polygon chain. (It may happen that a section contains points of multiple edges.)

For each refined section, then the farthest vertex and its shortest distance to the section can be calculated. Subsequently, δ can be calculated by a finite minimum according to (4.8).

Calculation of the distinguished chords

Let S be a refined section which realizes the minimum (4.8), so with $\delta = d_S(y(S))$. The two following lemmata describe how to obtain from such a refined section the distinguished chords.

5.1. Lemma. *Let S be a refined section with $\delta = d_S(y(S))$.*

If $z^ \in \overline{S}$ is the nearest point for $y(S)$, hence $\delta = d_S(y(S)) = d(z^*, y(S))$ then the segment $[z^*, y(S)]$ is a distinguished chord of P .*

Proof. One has to show

$$(5.1) \quad d(z^*, y(S)) \geq d(z^*, q) \quad \forall q \in P.$$

For this it suffices to show

$$(5.2) \quad d(z^*, y(S)) = \max_{q \in P} d(z^*, q).$$

Setting in (4.1): $x = z^*$ implies $\mu(z^*) = d(z^*, y(S))$. By $\mu(z^*) = \max_{q \in P} d(z^*, q)$, this expresses the assertion. \square

The farthest vertex $y(S)$ from the *inner* points of an arbitrary (refined) section S is always uniquely determined. This does not hold necessarily for the two ends of S . Let, for instance, A be the starting point of S . Admittedly, $y(S)$ is also the farthest vertex from A , but A may owe further farthest vertices.

5.2. Lemma. *Let S be a refined section with $\delta = d_S(y(S))$.*

If, for an initial point a of S and a vertex $\tilde{a} \neq a$ one has

$$(5.3) \quad \delta = d(a, y(S)) = d(a, \tilde{a})$$

then $[a, \tilde{a}]$ is also a distinguished chord of P .

Proof. The segment $[a, y(S)]$ is a distinguished chord, by Lemma 5.1, so

$$(5.4) \quad \delta = d(a, y(S)) \geq d(a, q) \quad \forall q \in P.$$

The same inequality holds, by (5.3), if $d(a, y(S))$ is replaced by $d(a, \tilde{a})$. Thus, $[a, \tilde{a}]$ is a distinguished chord. \square

Under the hypotheses just made, a necessarily is a section point because $y(S)$ and \tilde{a} have the same distance from a , so a lies on the perpendicular bisector of the chord $[y(S), \tilde{a}]$.

Obviously, the algorithm can be shaped in such a way that all distinguished chords which occur by the Lemmata 5.1 and 5.2 are generated in finitely many steps. In fact, *all* existing distinguished bisectors can be obtained in this way:

5.3. Lemma. *Any distinguished chord $[p_0, q_0]$ of the polygon P arises by the assumptions of Lemma 5.1 or Lemma 5.2.*

Proof. At any rate, there exists a refined section S such that p_0 lies in S or is the initial point of S . By assumption one has

$$(5.5) \quad \delta = d(p_0, q_0) \geq d(p_0, q) \quad \forall q \in P.$$

In particular, q_0 is a farthest vertex from p_0 . The following cases may occur:

(i) q_0 is also a farthest vertex for the points of S : Then $q_0 = y(S)$, because such a farthest vertex is uniquely determined. First, one has to show:

$$(5.6) \quad d_T(y(T)) \geq d_S(y(S)) \quad \text{for all refined segments } T :$$

This follows from

$$(5.7) \quad d_T(y(T)) \geq \delta = d(p_0, q_0) = d(p_0, y(S)) \geq d_{\bar{S}}(y(S)) = d_S(y(S)).$$

Specifically for $T = S$, here appears everywhere the equality sign: $d(p_0, q_0) = d_S(y(S))$. So p_0 is the nearest point for $y(S)$ in \bar{S} . Hence for $z^* := p_0$, all assumptions of Lemma 5.1 are fulfilled.

(ii) q_0 is *not* a farthest vertex for the points of S : Then $p_0 \notin S$, yet $p_0 =: a$ is the initial point of S . By the assumption (5.5) one has: $\delta(a, q_0) \geq d(a, q)$ for all $q \in P$. Thus, q_0 is a farthest point from a , but $q_0 \neq y(S)$.

One has $\delta = d(a, q_0) = d_S(y(S))$ because q_0 as well as $y(S)$ are farthest vertices from a . So, for S, a and $\tilde{a} := q_0$, the assumptions of Lemma 5.2 are satisfied. \square

Lemma 5.3 justifies the algorithm which rests on the Lemmata 5.1 and 5.2.

In the following pictures, the endpoints of the sections will be marked by little crosses and the distinguished chords as dashed segments.

6. Short polygons ($n \in \{3, 4\}$)

The general polygon case with n edges is completely open. Though many experiments seem to support the ‘isoperimetric inequality’ (2.1), we can prove it only for triangles and offer a concrete conjecture for quadruples, the latter with a surprising figure as a ‘magic quadrangle’ in the equality case.

Restricting oneself to the pure polygon case one may expect that the ‘isoperimetric quotient’ L/δ assumes a value bigger than π but with limit π for $n \rightarrow \infty$.

The case of triangles

6.1. Theorem. *For each triangle Δ in \mathbb{R}^2 one has*

$$(6.1) \quad L(\Delta) \geq 2\sqrt{3} \cdot \delta(\Delta),$$

with equality exactly for equilateral triangles.

The quotient L/δ is invariant under similarities. One may use this and other restrictions in order to build a ‘moduli space’ for all convex polygons of a fixed edge number n , thus reducing the problem to a study of L/δ on this moduli space. For triangles we shall follow a principle which is also sensible in higher cases, namely to single out a diameter and normalize it to a fixed size.

Diameter in the general polygon case

6.2. Lemma. *A diameter D of a convex polygon P always connects two vertices.*

Denote by H_1 and H_2 the two open half planes of the straight line spanned by D and by D° the relative interior of D .

If $P \cap H_i \neq \emptyset$ for an $i \in \{1, 2\}$ then $P \cap \overline{H_i} \setminus D^\circ$ lies schlicht over D in the direction orthogonal to D . If $P \cap H_i = \emptyset$ then $P \cap \overline{H_i}$ is an edge of P and equal to D .

Proof. The first part follows directly from Lemma 4.1 because each endpoint of a diameter is the farthest point from the other endpoint.

(*) Without restriction, let the endpoints of the diameter D be $A := (-1, 0)$ and $B := (1, 0)$ and assume that H_1 is the upper half plane and H_2 the lower half plane of \mathbb{R}^2 .

Each of the four edges through A or B makes an angle $< \pi/2$ with D because, otherwise, there resulted a contradiction to the maximal length of D . If one of these edges forms an angle 0 with D then D is an edge of P and P lies completely in \overline{H}_1 or in \overline{H}_2 .

If e.g. $P \cap H_1 \neq \emptyset$ then always one of the four edges through A, B is directed into H_1 (while the two others don't this). The part of P in \overline{H}_1 forms, together with D a convex polygon $P_1 := (P \cap \overline{H}_1) \cup D$ with D as an edge and diameter. Then the two other edges of P_1 through A, B make with D an angle between 0 and $\pi/2$. In this situation, a straight line orthogonal to D , $te_1 + \mathbb{R}e_2$, doesn't cut the polygon P_1 if $t < -1$ oder $t > 1$, and except in te_1 exactly once in a point $te_1 + f(t)e_2$ with $f(t) \geq 0$ if $-1 \leq t \leq 1$. Otherwise, in case $|t| > 1$, there were a contradiction to the diameter property of D . In case $|t| < 1$, the aforementioned straight line cannot have a whole segment in common with P_1 since, then, it were a support line of P_1 such that either A or B would not belong to P_1 . In case $t = \mp 1$, obviously A , resp. B is the sole common point of the straight line with P_1 . So the convex region C_1 bounded by P_1 is given by

$$(6.2) \quad C_1 = \{(t, s) \mid -1 \leq t \leq 1, 0 \leq s \leq f(t)\}$$

with a function $f : [-1, 1] \rightarrow \mathbb{R}$ with $f(-1) = f(1) = 0$ and $f(t) \geq 0$ for $|t| < 1$. In fact, f is a concave function since C_1 is convex and C_1 is the subgraph of f . By $f(-1) = f(1) = 0$ this implies $f(t) > 0$ for $|t| < 1$.

Projecting the edges of P_1 onto D one sees that the function f is piecewise affine.

If, e.g., $P \cap H_2 = \emptyset$ then $P \cap H_1 \neq \emptyset$ and the same arguments as above for P_1 apply for P itself.

We still record:

6.3. Remark. In case $P \cap H_1 \neq \emptyset$ there holds, under the normalization (*):

$$(6.3) \quad P \cap \overline{H}_1 = \{(t, f(t)) \mid -1 \leq t \leq 1\} \cup D,$$

with a concave and piecewise affine function $f : [-1, 1] \rightarrow \mathbb{R}$, where $f(-1) = f(1) = 0$ and $f(t) > 0$ for $|t| < 1$. \square

Proof of Theorem 6.1

In the triangle case things shall be arranged in the above manner. A diameter is necessarily one of the longest edges, so the segment from $A = (-1, 0)$ to $B = (1, 0)$ is an edge of maximal length. The edge of the middle size may start in A , the shortest edge in B . The third vertex $C = (x, y)$ then lies in the quadrant $\{(x, y) \mid x \geq 0, y \geq 0\}$. The moduli space sounds here

$$(6.4) \quad \mathcal{M} := \{(x, y) \mid x \geq 0, y \geq 0, (x+1)^2 + y^2 \leq 4\},$$

where, admittedly, the segment from $(0, 0)$ to B ($\iff y = 0$) corresponds to degenerate triangles ($\Omega = \emptyset$). In the sequel, it is assumed that $y > 0$.

The centers of the edges AB , BC , CA are

$$(6.5) \quad M := (0, 0), \quad M_+ := \frac{1}{2}(x+1, y), \quad M_- := \frac{1}{2}(x-1, y).$$

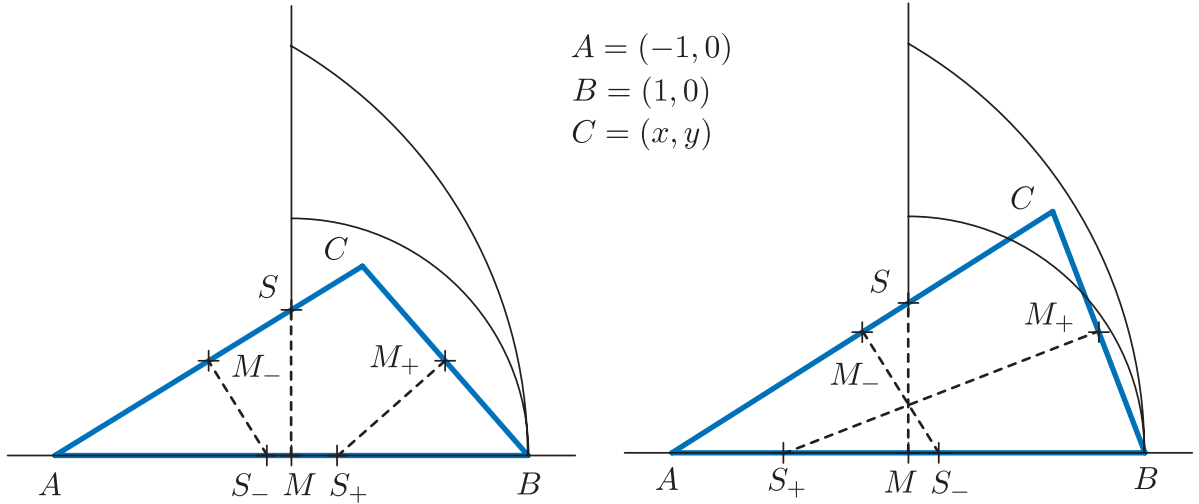
The section points of the perpendicular bisectors through M , M_+ , M_- with P are

$$(6.6) \quad S = (0, h), \quad S_+ = (x_+, 0), \quad S_- = (x_-, 0),$$

with

$$(6.7) \quad h := \frac{y}{1+x}, \quad x_+ := \frac{1}{2} \frac{1 - (x^2 + y^2)}{1 - x}, \quad x_- := \frac{1}{2} \frac{(x^2 + y^2) - 1}{1 + x}.$$

From this one reads that M_+ always lies right of the y -axis and M_- left of it and, further, that S_+ and S_- are situated on the same resp. different sides of the y -axis according to $x^2 + y^2 < 1$ resp. $x^2 + y^2 > 1$. Also, the angle at C follows these inequalities: In the first case it is obtuse, in the second it is sharp. This results from the calculation of the scalar product $\langle A - C, B - C \rangle = x^2 + y^2 - 1$. Cf. the following pictures:



In case $x^2 + y^2 = 1$ (right angle at C) nothing has to be changed w.r.t. M_+ and M_- , but S_+ and S_- coincide with M .

In the sequel we will refer to this situation as far as it is geometrically evident.

One realizes that the unit circle decomposes the moduli space \mathcal{M} into two parts which behave differently. In fact, the differences will become even more drastic with regard to the invariant δ .

Subcase $x^2 + y^2 \leq 1$

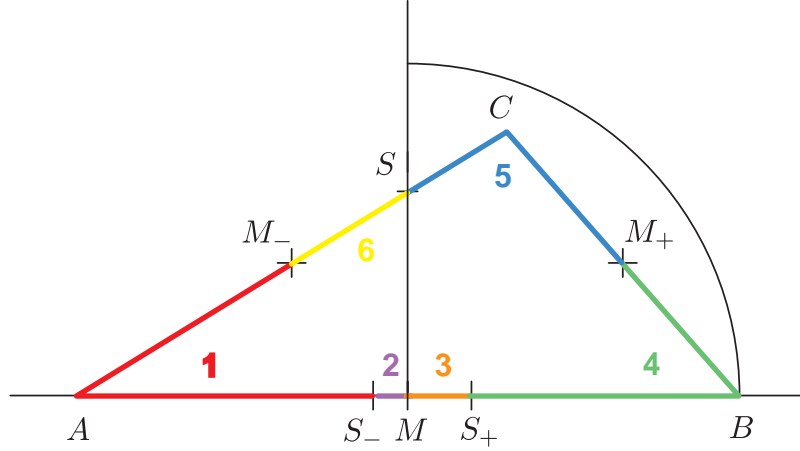
Under this assumption the behaviour is pretty simple and also uniform.

6.4. Lemma. *If, for $C \in \mathcal{M}$, one has $x^2 + y^2 \leq 1$ then $\delta = 1$, hence by $L > 4$:*

$$\frac{L}{\delta} > 4.$$

For those $C \in \mathcal{M}$ the inequality (6.1) is strict, so nothing has to be added.

Proof of 6.4. First of all, the case $x^2 + y^2 < 1$ will be treated explicitly. The decomposition of P into sections looks as follows:



In particular, some of the sections ‘round the corner’ (namely those with numbers 1, 4, and 5). This means that they contain vertices which are not section points themselves. If the vertices would be added to the decomposition then from the six sections there would arise nine refined sections. However, these nine sections have no impact on the following discussion.

For short, write $y_k := y(S_k)$ and $d_k := d_{S_k}(y(S_k))$, and denote by z_k^* a nearest point to y_k on S_k . These objects are determined according to the following table:

(6.8)

S_k	y_k	z_k^*	d_k
S_1	B	M_- or S_-	> 1
S_2	B	M	1
S_3	A	M	1
S_4	A	M_+ or S_+	> 1
S_5	A	M_+ or S	> 1
S_6	B	M_- or S	> 1

Most entries are immediately obvious. For example, for the section S_1 , one has $y_1 = B$ because $A \in S_1$ and also z_1^* is equal to M_- or to S_- , so in any case $d_1 > 1$. For the section S_2 one has $y_2 = B$, as shown by comparison with $d(M, C) = 1$, and evidently $z_2^* = M$, hence $d_2 = 1$. The following lines in the table are similarly simple. In some cases it is of some importance that the orthogonal footpoints of A and B on the opposite edges come to lie outside of these.¹

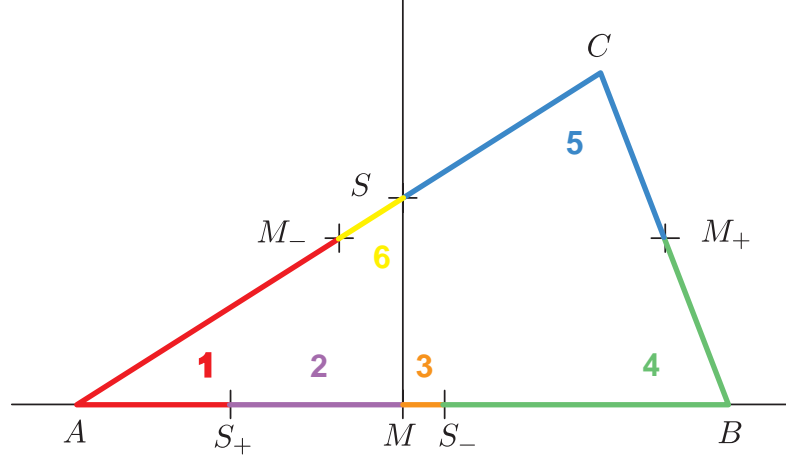
By this, the minimal d_k is 1, thus $\delta = 1$. Distinguished chords by Lemma 5.1 are those from M to A and to B . To be added is another distinguished chord by Lemma 5.2 from M to C .

¹A more detailed determination of z_k^* in the ‘or-cases’ of the table (6.8) is possible but not really necessary here.

The case $x^2 + y^2 = 1$ resolves itself in the same manner or, also, by performing the limit $x^2 + y^2 \uparrow 1$, applying Lemma 3.1, part (i). Namely, a triangle with $x^2 + y^2 = 1$ can be obtained in a continuous manner from a triangle with $x^2 + y^2 < 1$, e.g. by a radial deformation with center 0. Hereby, the relation between the triangles can be kept bijectively and uniformly small.

Subcase $x^2 + y^2 > 1$

Here, more possibilities are to be considered, and the discussion is clearly more complicated. The decomposition of P looks as follows:



In order that the sections look really that way (i.e. that the five points on $[A, B]$ and the four points on $[A, C]$ are pairwise distinct) it is necessary to assume:

$$(6.9) \quad x > 0, \quad (x+1)^2 + y^2 < 4.$$

One can confirm this constellation with the formulae (6.5) to (6.7). By (6.9), the two equilateral cases $d(A, S) = d(B, S)$ resp. $d(A, C) = d(A, B) = 2$ are excluded (they will be discussed later on).

The difference to the preceding case is the interchange in the position of S_+ und S_- and the fact that a sharp angle now arises at C .

It is suitable to calculate first the farthest points of the sections. Hereby, one can save some effort by regarding the following principle on the role of the perpendicular bisectors as discussed in section 4:

6.5. Lemma. *If a point $X \in P$ moves strictly monotonously towards a point which lies on exactly one perpendicular bisector (of two vertices P, Q) and if the farthest vertex from X has been P before that it is equal to Q afterwards. \square*

It is immediately clear that

$$(6.10) \quad y_1 = B, \quad y_4 = A, \quad y_5 = A,$$

since the sections with numbers 1, 4, 5 each contain a vertex. With the principle 6.5 this implies

$$(6.11) \quad y_2 = C, \quad y_3 = C, \quad y_6 = B.$$

Now the distances d_k , i.e. the footpoints z_k^* of y_k on \bar{S}_k must be determined.

For z_1^* :

The footpoint of $y_1 = B$ on the straight line $A \vee C$ lies outside the segment $[A, M_-]$ because the angle $\sphericalangle(AM_-B)$ is obtuse. This follows from a calculation which relies on Eqns. (6.5) to (6.7):

$$(6.12) \quad \langle A - M_-, B - M_- \rangle = \frac{1}{4}((x-1)^2 + y^2 - 4) < 0.$$

The last estimate is valid since otherwise one had $x < 0$, using (6.9).

Thus the foot of B on \bar{S}_1 equals M_- or S_+ . The decision for this results from the sign of the quantity

$$(6.13) \quad \varphi_1(x, y) := 4 \cdot (|M_- - B|^2 - |S_+ - B|^2) = (1-x)^2((x-3)^2 + y^2) - ((x-1)^2 + y^2)^2.$$

For $x \in [0, 1]$, the sign condition for $\varphi_1(x, y)$ can be rearranged by solution for y . This yields

$$(6.14) \quad z_1^* = \begin{cases} M_- & \text{for } y \geq \psi_1(x) \\ S_+ & \text{for } y \leq \psi_1(x), \end{cases}$$

where

$$(6.15) \quad \psi_1(x) := \sqrt{\frac{1-x}{2}} \sqrt{\sqrt{(9-x)^2 - 48} - (1-x)}.$$

In the equality case of (6.14), both points M_- and S_+ are footpoints of y_1 on S_1 (with same distances).

The graph of the function ψ_1 over $[0, 1[$ lies there between the circles $\mathbb{S}(M, 1)$ and $\mathbb{S}(A, 2)$ (with equality at $x = 1$); for one confirms that

$$(6.16) \quad \sqrt{1-x^2} < \psi_1(x) < \sqrt{4-(1+x)^2} \quad \text{for } x \in [0, 1[.$$

For z_2^* :

Obviously $z_2^* = M$ (consider the circle through M around $y_2 = C$).

To z_3^* :

The footpoint of $y_3 = C$ on S_3 is equal to S_- or to the footpoint $(x, 0)$ of C on the straight line $A \vee B$, according to the sign of $x - x_-$. One calculates

$$(6.17) \quad x - x_- = \frac{(1+x)^2 - y^2}{2(1+x)},$$

hence

$$(6.18) \quad z_3^* = \begin{cases} S_- & \text{for } y \geq \psi_3(x) \\ (x, 0) & \text{for } y \leq \psi_3(x), \end{cases} \quad \psi_3(x) := 1 + x.$$

For z_4^* :

As z_4^* , only the points M_+ , S_- come in consideration. The foot of $y_4 = A$ on the segment $[M_+, B]$ is really equal to M_+ because the angle $\sphericalangle(AM_+B)$ is obtuse. This follows from the calculation

$$(6.19) \quad \langle A - M_+, B - M_+ \rangle = \frac{1}{4}((1+x)^2 + y^2 - 4) < 0.$$

In fact, the point S_- is always closer to S_4 as M_+ . This rests on the inequalities

$$(6.20) \quad d(A, M_+) > 1 + \frac{1}{2}(1+x) > x_-.$$

The first inequality expresses the fact that the orthogonal projection of the segment $[A, M_+]$ onto the x -axis shortens the segment. The second inequality reduces itself to $y < \sqrt{4+4x}$ by conversion, and this is certainly true for $y \leq \sqrt{4-(1+x)^2}$. Thus:

$$(6.21) \quad z_4^* = S_-.$$

For z_5^* :

Similar to the case z_4^* , the footpoint L_5 of $y_5 = A$ on the straight line $B \vee C$ now lies on the segment $[M_+, C]$. In the rectangular triangle AL_5C , the leg $[A, L_5]$ is shorter than the hypotenuse $[A, C]$, so

$$(6.22) \quad z_5^* = S.$$

For z_6^* :

By (6.12), the angle $\sphericalangle(BM_-S)$ is sharp, hence the footpoint of $y_6 = B$ on S_6 is equal to S or to the footpoint L_6 of B on the straight line $C \vee A$. Decisive for this is the scalar product

$$(6.23) \quad \langle A - S, B - S \rangle = h^2 - 1 = \frac{(y - (1+x))(y + 1+x)}{(1+x)^2}.$$

If the scalar product is ≤ 0 then $z_6^* = S$; if it is ≥ 0 then $z_6^* = L_6$ where by (4.7):

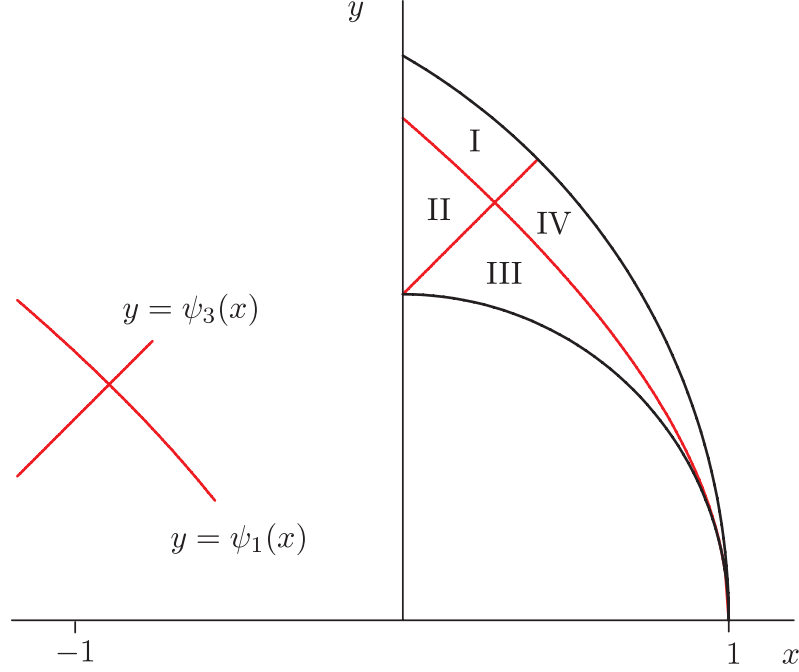
$$(6.24) \quad L_6 = \frac{\langle B - C, A - C \rangle A - \langle B - A, A - C \rangle C}{|A - C|^2}.$$

The distance $d(B, S)$ resp. $d(B, L_6)$ itself is calculated as follows, using (4.5)

$$(6.25) \quad d_6 := \begin{cases} \frac{2y}{\sqrt{(1+x)^2 + y^2}} & \text{for } y \geq \psi_3(x) \\ \frac{\sqrt{(1+x)^2 + y^2}}{1+x} & \text{for } y \leq \psi_3(x) \end{cases}$$

with ψ_3 as in (6.18).

The region which is now of interest lies between the circle arcs $\mathbb{S}(M, 1)$ and $\mathbb{S}(A, 2)$ (always in the first quadrant) and is divided into four subregions I, II, III, IV, according to the cases (6.14) and (6.18). Cf. the following picture



The key at left indicates the equations describing the separating curves. The following features will justify this appearance.

The function ψ_1 has the boundary values $\psi_1(0) = \sqrt{\sqrt{33} - 1}/\sqrt{2} = 1.5402\dots$, $\psi_1(1) = 0$ and it decreases strictly monotonously in between. Namely, the derivative of ψ_1 on $]0, 1[$ equals, up to a positive factor, the following expression which can be easily estimated:

$$(6.26) \quad (1 - x) \cdot \sqrt{(x - 9)^2 - 48} - ((x - 7)^2 - 28) < 1 \cdot \sqrt{16} - 21 < 0.$$

The graph of the function ψ_3 over $[0, 1]$ cuts the circle $\mathbb{S}(A, 2)$ at the point $(\sqrt{2} - 1, \sqrt{2}) = (0.4142\dots, 1.4142\dots)$. Since $\psi_1(\sqrt{2} - 1) = \sqrt{8\sqrt{2} - 10}$, the two graphs of ψ_1 and ψ_3 over $[0, \sqrt{2} - 1]$ intersect themselves exactly once. The condition for this intersection leads to the cubic equation $-3 + 12x - 6x^2 + 4x^3 + x^4 = 0$ which has a unique real solution in $[0, 1]$ (computed numerically as) $0.2817\dots$. So, the section point sounds $(0.2817\dots, 1.2817\dots)$.

Now, for each region, we must deduce the six distances $d_k = d(y_k, z_k^*)$ from the above information on the foci z_k^* and also their minima. The following table contains the results

for the distances:

$$(6.27) \quad \begin{array}{c|c|c|c|c} d_k & \text{region I} & \text{region II} & \text{region III} & \text{region IV} \\ \hline d_1 & \frac{1}{2}\sqrt{(3-x)^2+y^2} & \frac{1}{2}\frac{(1-x)^2+y^2}{1-x} & \frac{1}{2}\frac{(1-x)^2+y^2}{1-x} & \frac{1}{2}\sqrt{(3-x)^2+y^2} \\ d_2 & d_2 & d_2 & d_2 & d_2 \\ d_3 & y & y & d_4 & d_4 \\ d_4 & d_4 & d_4 & d_4 & d_4 \\ d_6 & \frac{2y}{\sqrt{(1+x)^2+y^2}} & \frac{2y}{\sqrt{(1+x)^2+y^2}} & & \end{array}$$

where

$$(6.28) \quad d_2 = \sqrt{x^2+y^2}, \quad d_4 = \frac{1}{2} \frac{(1+x)^2+y^2}{1+x}, \quad d_5 = \frac{\sqrt{(1+x)^2+y^2}}{1+x}.$$

Obviously

$$d_5 = \sqrt{\frac{2}{1+x}} \sqrt{d_4} > d_4$$

with equality on $\mathbb{S}(A, 2)$. So, for the search of the minima, the d_5 can be ignored. Therefore the line for d_5 is omitted in the table, and similar for the two entries downright (which also contained d_5).

Intermezzo

In some cases, the handling of the inequalities is facilitated by passing from the cartesian coordinates x, y to the elliptic coordinates u, v defined by

$$(6.29) \quad \begin{aligned} u &= \frac{1}{2} \left(\sqrt{(1+x)^2+y^2} + \sqrt{(1-x)^2+y^2} \right) & x &= uv \\ v &= \frac{1}{2} \left(\sqrt{(1+x)^2+y^2} - \sqrt{(1-x)^2+y^2} \right) & y &= \sqrt{(u^2-1)(1-v^2)} \end{aligned}$$

The equations for u, v are built from the distances to the points A, B . They will only be considered in the upper half plane ($\{(x, y) \mid y \geq 0\}$). By the equations (6.29) the upper half plane is homeomorphically mapped onto the region $\{(u, v) \mid u \geq 1, -1 \leq v \leq 1\}$. The transformation back is recorded right in (6.29). Outside the x -axis the map $(x, y) \mapsto (u, v)$ is diffeomorphic of class C^∞ .

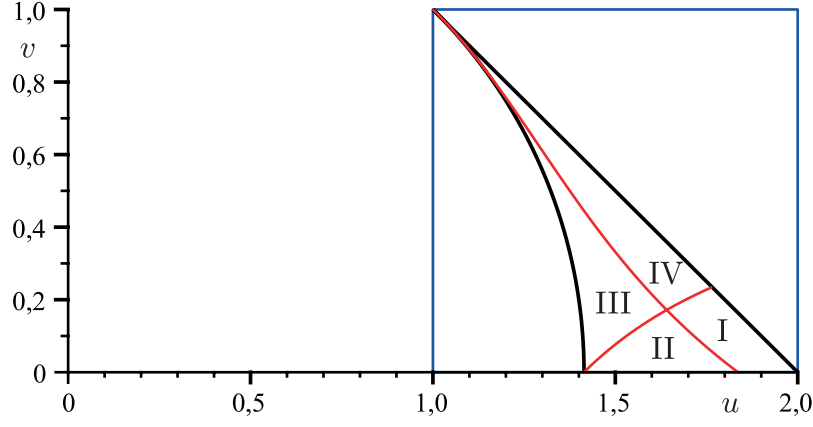
If the fundamental region in the first quadrant between $\mathbb{S}(M, 1)$ and $\mathbb{S}(A, 2)$ is transformed with the elliptic coordinates then the image lies in the square $\{(u, v) \mid 1 \leq u \leq 2, 0 \leq v \leq 1\}$ and is given there by

$$(6.30) \quad u^2 + v^2 \geq 2, \quad u + v \leq 2.$$

The curves $y = \psi_1(x)$, $y = \psi_2(x)$ which separate the fundamental region into four parts are subsequently given by

$$(6.31) \quad (1-uv)^2((u-v)^2+4(2-uv)) - (u-v)^4 = 0, \quad (u^2-1)(1-v^2) - (1+uv)^2 = 0$$

with the corresponding assignment of the boundaries:



Moduli space in elliptic coordinates

The distinctive points on the u -axis have the coordinates

$$(6.32) \quad \sqrt{2} = 1.4142\dots, \quad \sqrt{\frac{\sqrt{33}+1}{2}} = 1.8363\dots$$

The section point of the separating lines and the section point on the line $u + v = 2$ are

$$(6.33) \quad (1.6409\dots, 0.1716\dots), \quad \left(1 + \sqrt{2 - \sqrt{2}}, 1 - \sqrt{2 - \sqrt{2}}\right) = (1.7653\dots, 0.2346\dots).$$

End of Intermezzo

All the regions will now be examined under the conditions

$$(6.34) \quad x > 0, \quad y > 0, \quad x^2 + y^2 > 1, \quad (1+x)^2 + y^2 < 4.$$

The regions I to IV are meant *without* these points, however with inclusion of the separating curves $y = \psi_1(x)$ and $y = \psi_2(x)$. The comparisons of the d_k among themselves are done including the equality signs. The names I to IV will be maintained for the image regions in elliptic coordinates (using the corresponding relations of the boundaries).

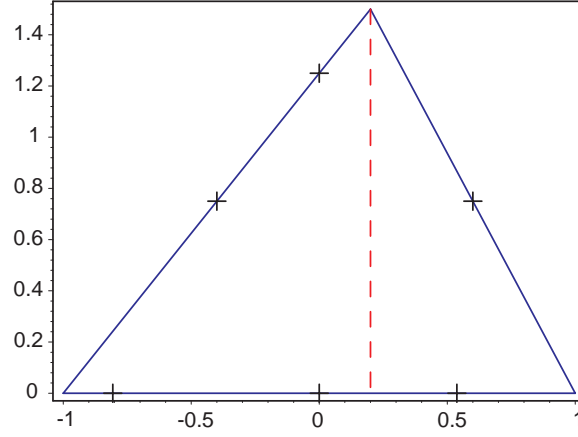
For region I:

Under the additional conditions

$$(6.35) \quad y \geq \psi_1(x), \quad y \geq \psi_3(x)$$

one must find the minimum of the entries in the column ‘region I’ of Table (6.27).

The following picture shows a typical pattern for this region



Hereby, one may conjecture that the minimum is $d_3 = y$. This is really confirmed by the following comparisons:

Comparison with d_1 : The inequality $d_3 \leq d_1$ reduces itself to

$$(6.36) \quad y \leq \sqrt{3} - \frac{x}{\sqrt{3}}.$$

This inequality describes the ordinate set of the tangent of $\mathbb{S}(A, 2)$ at $(0, \sqrt{3})$, so is satisfied in all regions.

Comparison with d_2 : This is obvious (equality for $x = 0$).

Comparison with d_4 : The relation $d_3 \leq d_4$ is the inequality for the arithmetic and geometric mean for $(1+x)^2$ and y^2 .

Comparison with d_6 : This is obvious.

The strict inequality (6.1) says in the present case:

$$(6.37) \quad 2 + \sqrt{(1+x)^2 + y^2} + \sqrt{(1-x)^2 + y^2} > 2\sqrt{3}y$$

or, in the elliptic coordinates

$$(6.38) \quad 1 + u > \sqrt{3} \sqrt{(u^2 - 1)(1 - v^2)}.$$

After squaring, this is equivalent to

$$(6.39) \quad (u + 1)(3uv^2 - 3v^2 - 2u + 4) > 0.$$

By $u + 1 \geq 2$ this is reduced to

$$(6.40) \quad 3(u - 1)v^2 > 2(u - 2).$$

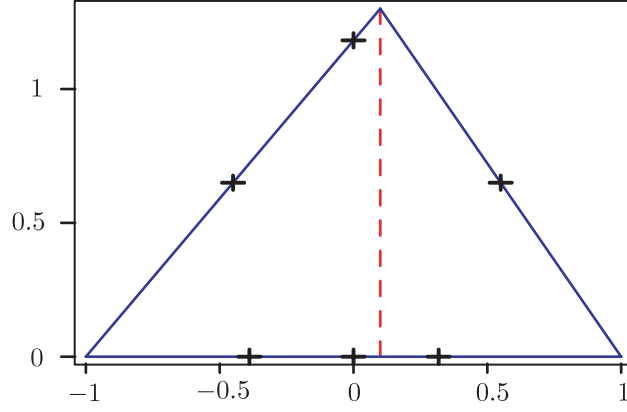
But this inequality is valid by observing $v > 0$ and $1 < u \leq 2$.

For region II:

Under the additional conditions

$$(6.41) \quad \psi_3(x) \leq y \leq \psi_1(x)$$

one must find the minimum of the entries in the column ‘region II’ of Table (6.27).



This picture shows a typical pattern for this region, and again, one can conjecture that the minimum is $d_3 = y$. This is confirmed by the following comparisons.

Comparison with d_1 : The inequality $d_3 \leq d_1$ reduces itself to the arithmetic-geometric inequality for $(1-x)^2$ and y^2 .

Comparison with d_2, d_4, d_6 : This runs as for region I.

The strict inequality (6.1) is here expressed in the same way as for the region I, i.e. by (6.37) and, equivalently, by (6.40). Here, one has $2(u-2) < 2(1.9-2) = -0.2$ by (6.32), hence $3(u-1)v^2 > 2(u-2)$.

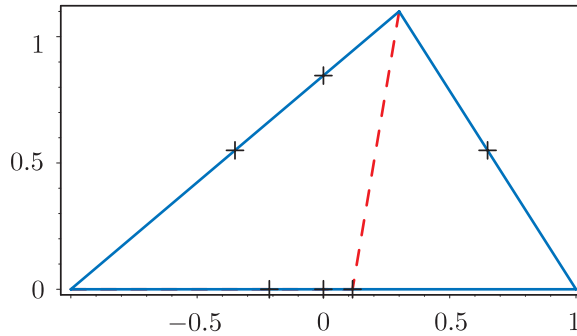
For region III:

Under the additional conditions

$$y \leq \psi_1(x), \quad y \leq \psi_3(x)$$

one must find the minimum of the entries in the column ‘region III’ of Table (6.27).

The following picture shows a typical pattern for this region:



Again, one can conjecture that the minimum is d_4 . This is confirmed by the following comparisons:

Comparison with d_1 :

The inequality $d_4 \leq d_1$ is equivalent to

$$(6.42) \quad 2x(x^2 + y^2 - 1) \geq 0.$$

Obviously, this is true.

Comparison with d_2 :

The inequality $d_4 \leq d_2$ is equivalent to

$$(6.43) \quad (x^2 + y^2 - 1)(3x^2 + 4x + 1 - y^2) \geq 0.$$

The first factor is positive, the second factor is ≥ 0 if

$$(6.44) \quad y \leq \sqrt{(1+x)(3x+1)}.$$

By $y \leq \psi_3(x) = 1+x$ this is certainly true if

$$1+x \leq \sqrt{(1+x)(3x+1)}.$$

After squaring and cancelling of $1+x > 0$ this is equivalent to $x \geq 0$. So, (6.44) and then (6.43) are satisfied.

The assertion (6.1) now says in the strict case:

$$(6.45) \quad 2 + \sqrt{(1+x)^2 + y^2} + \sqrt{(1-x)^2 + y^2} > \sqrt{3} \frac{(1+x)^2 + y^2}{1+x}$$

or, equivalently with elliptic coordinates

$$(6.46) \quad 1+u > \frac{\sqrt{3}}{2} \frac{(u+v)^2}{1+uv}.$$

Since, in the region III, the y -coordinate is always $\leq \sqrt{2}$, it is sufficient for the proof of (6.46) to assume $y \leq y_0 := \sqrt{2}$, what means in elliptic coordinates $(u^2 - 1)(1 - v^2) \leq 2$ or

$$(6.47) \quad v^2 \geq 1 - \frac{2}{u^2 - 1} u^2 - 1.$$

As long as the right hand side is ≤ 0 (including $= -\infty$) this is trivially true. This right hand side is strictly monotonously increasing as a function of u and has the unique zero $\sqrt{3}$. So, one must show: From

$$(6.48) \quad u \geq \sqrt{3}, \quad v \geq \sqrt{1 - \frac{2}{u^2 - 1}} =: g(u)$$

follows, in the region III: (6.46), i.e.

$$(6.49) \quad (1+u)(1+uv) - \frac{\sqrt{3}}{2}(u+v)^2 > 0.$$

This inequality can explicitly solved for v :

$$(6.50) \quad f(u) < v < F(u)$$

with

$$(6.51) \quad \begin{aligned} f(u) &:= \frac{u}{\sqrt{3}}(u - (\sqrt{3} - 1)) - \frac{u+1}{\sqrt{3}} \sqrt{(u - \sqrt{3})^2 + 2\sqrt{3} - 3} \\ F(u) &:= \frac{u}{\sqrt{3}}(u - (\sqrt{3} - 1)) + \frac{u+1}{\sqrt{3}} \sqrt{(u - \sqrt{3})^2 + 2\sqrt{3} - 3}. \end{aligned}$$

The function F is, for $u \in [1, 2]$, always > 1 . For this, first consider the equation $F(u) = 1$ for all $u \in \mathbb{R}$. The solutions of this equation result by equating the left hand side of (6.49) to zero if $v = 1$, so from $(u+1)^2(2-\sqrt{3}) = 0$, and this is only true for $u = -1$. By $F(2) = 1 + \sqrt{3}$ one has $F(u) > 1$ for $u \in [1, 2]$. So, the second inequality in (6.50) is automatically satisfied for $1 \leq u \leq 2$, $0 \leq v \leq 1$.

In $[1, 2]$, the function f changes its sign exactly twice. For this, consider the equation $f(u) = 0$ first for all $u \in \mathbb{R}$. The solutions of this equation result by equating the left hand side of (6.49) to zero if $v = 0$, so from $\sqrt{3}u^2 - 2u - 2 = 0$. The solutions are

$$(6.52) \quad u_0 := \frac{1}{\sqrt{3}} \left(1 + 2\sqrt{1 + 2\sqrt{3}} \right) = 1.7972 \dots, \quad \frac{1}{\sqrt{3}} \left(1 - 2\sqrt{1 + 2\sqrt{3}} \right) = -0.64249 \dots$$

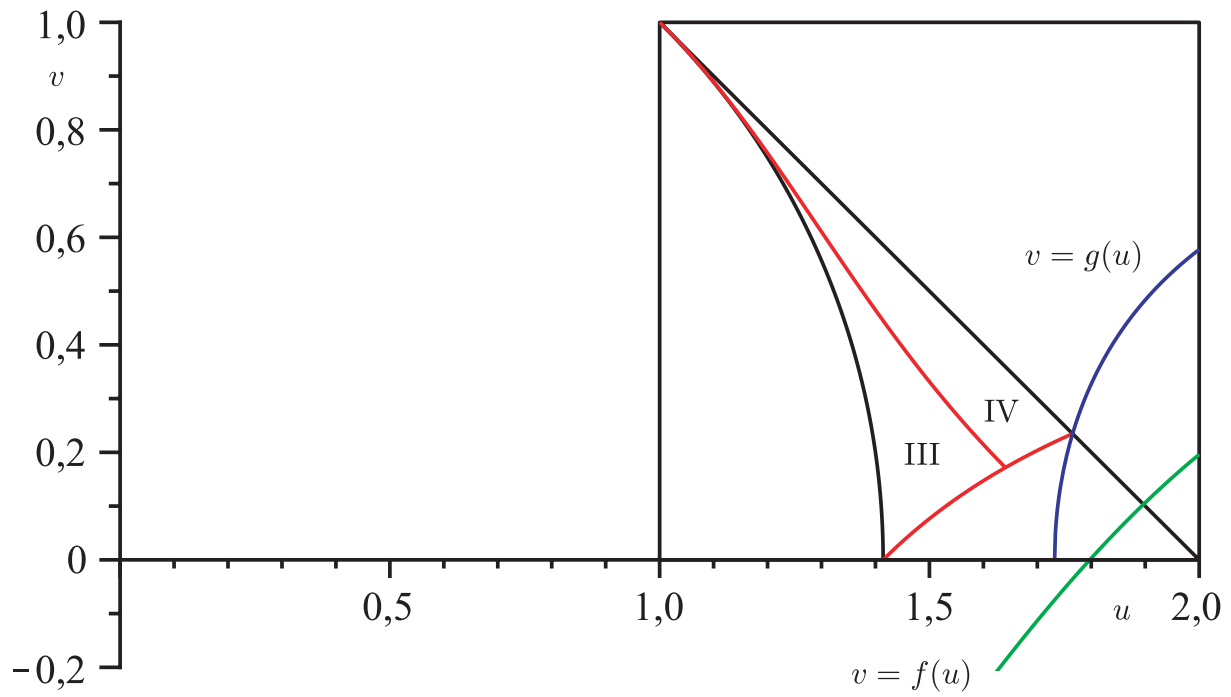
By $f(0) = -1$, one has $f(u) < 0$ for $u \in [1, u_0[$ and $f(u) > 0$ for $u \in]u_0, 2]$.

The graph of the function g from (6.48) cuts the line $u + v = 2$ at

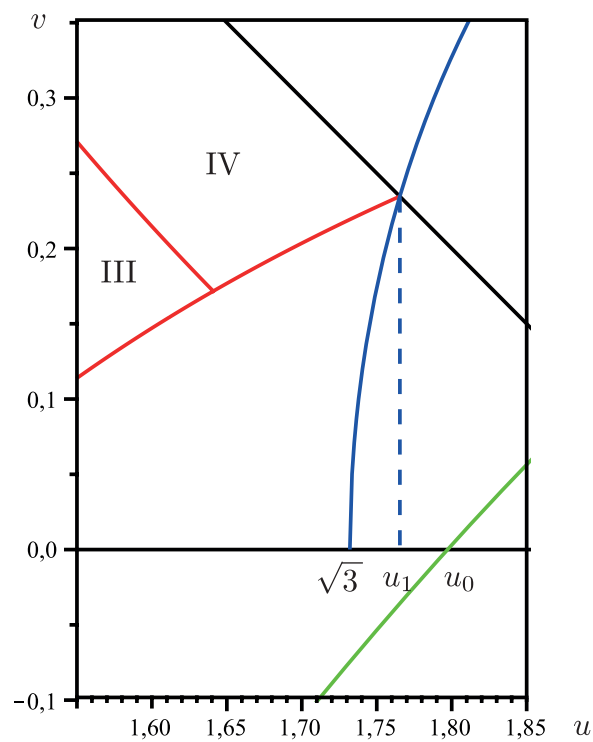
$$(6.53) \quad u = u_1 := 1 + \sqrt{2 - \sqrt{2}} = 1.7653 \dots$$

Thus, the points (u, v) with (6.48) only can belong to the transformed moduli space if $u \in [\sqrt{3}, u_1]$. But in this interval, one has $f(u) < 0$ (since $u_1 < u_0$) so (6.50) is satisfied there. Hereby it is evident for the region III that the asserted inequality is valid in a strict form.

The situation is elucidated by the following picture:



and again under magnification around the points $\sqrt{3}, u_1, u_0$ on the u -axis:



Magnification of the critical part

In a moment, it will come out that the same picture is also appropriate for the last region IV.

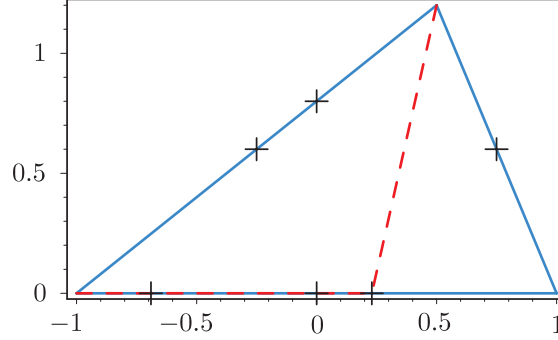
For region IV:

Under the additional conditions

$$\psi_1(x) \leq y \leq \psi_3(x)$$

one must find the minimum of the entries in the column ‘region IV’ of Table (6.27).

The following picture shows a typical pattern for this region.



One may conjecture that again d_4 yields the minimum. This will be confirmed by the following comparisons:

Comparison with d_1 : The inequality $d_4 \leq d_1$ is equivalent to

$$(6.54) \quad ((1+x)^2 + y^2)^2 - (1+x)^2((x-3)^2 + y^2) \leq 0.$$

This is equivalent to

$$(6.55) \quad y^2 \leq \frac{1+x}{2} \left(\sqrt{x^2 - 30x + 33} - (1+x) \right).$$

The expression in the large parentheses on the right turns out to be ≥ 0 such that one has to confirm that in the region IV the following holds:

$$(6.56) \quad y \leq \sqrt{\frac{1+x}{2}} \sqrt{\sqrt{x^2 - 30x + 33} - (1+x)} =: h(x).$$

The graph of h over $[0, 1]$ mostly runs outside the moduli space. One has

$$(6.57) \quad h(0) = \sqrt{\frac{\sqrt{33}-1}{2}} < \sqrt{3},$$

so $(0, h(0)) \in \mathcal{M}$, but this graph cuts $\mathbb{S}(A, 2)$ for $0 < x < 1$ exactly once namely for

$$(6.58) \quad x = x_0 := \frac{\sqrt{17}-3}{4} = 0.2807 \dots$$

This results by solving the equation $h(x) = \sqrt{4 - (1+x)^2}$ which is equivalent to $(1-x)(2x^2 + 3x - 1) = 0$. So, (6.56) is satisfied for all $x \in \mathcal{M}$ with $x \geq x_0$. Since all points (x, y) in the

closure of the region IV have an abscissa $x \geq 0.2817 \dots$ it is evident that Eqn. (6.54) is valid there always.

Comparison with d_2 : This has already been confirmed with region III.

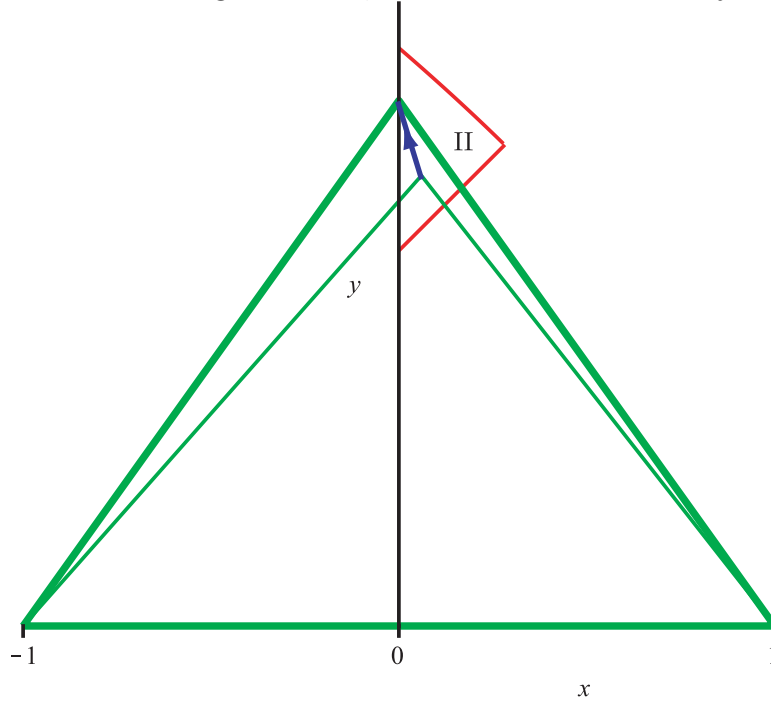
Further comparisons are not necessary, cf. Eqn. (6.27).

As for region III it remains to prove the inequality (6.45). However, this runs exactly by the same arguments as for region III because the bound $y_0 = \sqrt{2}$ there has been chosen in such a way that all points in the closure of IV also have an ordinate $y \leq y_0$.

Finally, the cases excluded with (6.34) must be examined. For this one still has $x^2 + y^2 > 1$.

For $x = 0$, where $1 < y \leq \sqrt{3}$:

Such a point (x, y) can be obtained by a limit process in the sense of Lemma 3.1, part (i), of points in the interior of the region I or II, with the effect that $\delta = y$ holds true:



In both cases one must show

$$(6.59) \quad 2 + 2\sqrt{1 + y^2} \geq 2\sqrt{3}y$$

with equality exactly for $y = \sqrt{3}$.

Equivalent with this is, by rearranging, $0 \geq 2y(y - \sqrt{3})$. This is obvious with equality exactly for $y = \sqrt{3}$.

For $(x, y) \in \mathbb{S}(A, 2)$, where $0 \leq x < \sqrt{2} - 1$:

Again, such a point (x, y) can be obtained as a limit of points in the interior of the region I, using Lemma 3.1, part (i). Regarding $\delta = y$ and $(1 + x)^2 + y^2 = 4$, one has to show:

$$(6.60) \quad 2 + 2 + \sqrt{(1 - x)^2 + 4 - (1 + x)^2} \geq 2\sqrt{3}y$$

with equality exactly for $(x, y) = (0, \sqrt{3})$. Rearranging (6.60) yields

$$(6.61) \quad \frac{1}{16}x(x+3)(8-3x-9x^2) \geq 0.$$

The zeros of the last polynomial factor are

$$(6.62) \quad \frac{-1 - \sqrt{33}}{6} = -1.1240\dots, \quad \frac{-1 + \sqrt{33}}{6} = 0.7907\dots$$

So, this factor is positive for $0 \leq x < \sqrt{2} - 1 = 0.4142\dots$, and Eqn. (6.61) is fulfilled with equality exactly for $x = 0$ and thus $y = \sqrt{3}$.

For $(x, y) \in \mathbb{S}(A, 2)$, where $\sqrt{2} - 1 \leq x < 1$:

Such a point (x, y) can be obtained as a limit of points in the interior of the region IV such that $\delta = d_4$ holds, again using Lemma 3.1, part (i). By $(1+x)^2 + y^2 = 4$ one has to show

$$(6.63) \quad 2 + 2 + \sqrt{4 - 4x} > \sqrt{3} \frac{4}{1+x},$$

equivalently

$$(6.64) \quad \sqrt{1-x} > \frac{2(\sqrt{3}-1) - 2x}{1+x}.$$

Here, the right hand side is negative for $x > \sqrt{3} - 1$ such that Eqn. (6.64) has only to be confirmed for $\sqrt{2} - 1 \leq x \leq \sqrt{3} - 1$. In this case, the assertion (6.64) turns out to be equivalent to $0 > x^2 - 2x + 2\sqrt{3} - 3$. The last polynomial has the zeros

$$(6.65) \quad 2 - \sqrt{3} = 0.2679\dots, \quad \sqrt{3} = 1.7320\dots,$$

so is negative in between, in particular for $x \in [\sqrt{2} - 1, \sqrt{3} - 1] = [0.3142\dots, 0.7320\dots]$.

So everything has been shown. \square

The case of quadrangles

Here, we only can offer a strong conjecture which causes a certain surprise because the extreme figure is definitely not the square.

6.6. Example. For a *square* P one has

$$(6.66) \quad \frac{L(P)}{\delta(P)} = \frac{8}{5}\sqrt{5} = 3.577\,708\,763\dots$$

See the following picture.

For a *general quadrangle* P we may choose the line from $B = (-1, 0)$ to $E = (1, 0)$ as a diameter D (see Lemma 6.2).

If D is not an edge of P we may place one more vertex as $A = (u, u)$ with $-1 < u < 1$, $u > 0$ in the upper half plane and the last vertex $C = (v, -v)$ with $-1 < v < 1$, $v > 0$ in

the lower half plane. Then a simple search programme shows an approximate minimum of the quotient $L(P)/\delta(P)$ at

$$u_0 = v_0 = 0, \quad u := 1.472, \quad v = 0.217$$

So, likely, the minimum figure is symmetric w.r.t. the y -axis. Moreover, the algorithm from above shows that the distinguished chords all have the same length (and are grounding orthogonally on the receiving edges).

Now, an exact search can be started, considering only the fixed vertices A and E and the variable vertices $A = (0, u)$ and $C = (0, -v)$, thus reducing the problem from dimension 4 to dimension 2, with $(u, v) \in \mathbb{R}^+ \times \mathbb{R}^+$ and the side condition $u \neq v$ ($u = v$ would yield a square). Also, the diameter of length 2 between A and E results in some additional restrictions.

The equality of the two perpendiculars from A to $B \vee C$, resp. from B to $A \vee E$ is expressed by the relation

$$\frac{1}{2} \frac{u+v}{\sqrt{v^2+1}} = \frac{u}{\sqrt{u^2+1}}$$

which comes down algebraically to

$$(u^3 - 3u + (3u^2 - 1)v)(u - v) = 0.$$

or, using $u - v \neq 0$, to

$$(6.67) \quad v = \frac{3 - u^2}{3u^2 - 1} u.$$

where $0 < u < 1$ and $v > 0$ require $u > 1/\sqrt{3}$.

So the problem is reduced to the minimum question of a real function, namely of the quotient ‘perimeter/length of the perpendicular from B to $A \vee E$ ’, i.e.

$$(6.68) \quad \frac{2\sqrt{u^2+1} + 2\sqrt{v^2+1}}{\frac{u}{\sqrt{u^2+1}}},$$

with v inserted according to (6.67). This function has the representation

$$f(u) := 4u \frac{u^2+1}{3u^2-1}, \quad \frac{1}{\sqrt{3}} < u < 1.$$

Its minimum condition is $3u^4 - 6u^2 - 1 = 0$, with the only positive solution

$$u = \frac{\sqrt{3}}{3} \sqrt{3 + 2\sqrt{3}}.$$

From this one derives the corresponding v with (6.67) and the isoperimetric quotient with (6.68), and one can state:

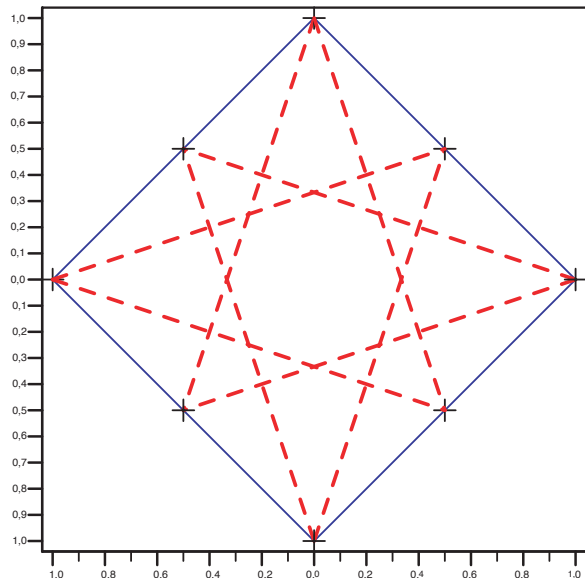
If P is the **kite** with the vertices

$$(6.69) \quad (-1, 0), \quad (1, 0), \quad \left(0, \frac{\sqrt{3}}{3} \sqrt{2\sqrt{3}+3}\right), \quad \left(0, -\frac{1}{3} \sqrt{2\sqrt{3}-3}\right),$$

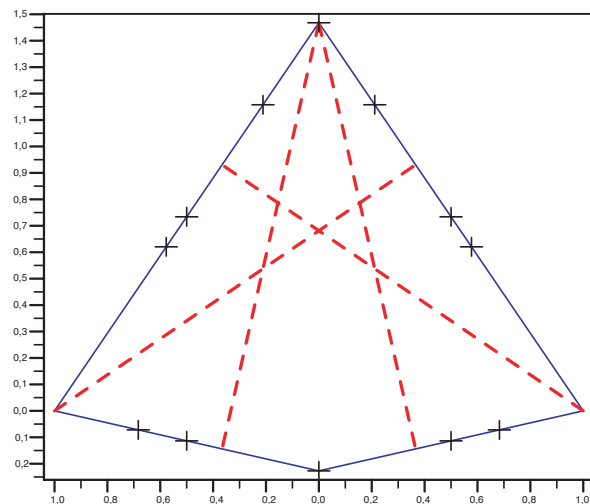
then

$$(6.70) \quad \frac{L(P)}{\delta(P)} = \frac{4}{3} \sqrt{2\sqrt{3} + 3} = 3.389\,946\,342\dots$$

Due to these weird number relations, this figure P shall be called a **magic kite**.



SQUARE



MAGIC KITE

each with distinguished chords (dashed)
and section points (crosses)

If D is an edge, the two other vertices A and C of P lie in the same half plane, say in the upper one. Numerical optimization programmes don't give any indication that, varying A and C , could yield a genuine convex quadrangle with minimal invariant δ .

These arguments provide a strong support for the

Conjecture 2: *For any convex quadrangle $P \subset \mathbb{R}^2$ there holds the inequality*

$$L(P) \geq \frac{4}{3} \sqrt{2\sqrt{3} + 3} \cdot \delta(P)$$

between the perimeter $L(P)$ and the invariant $\delta(P)$, with equality exactly for magic kites.

References

Amir, D., and Ziegler, Z. [1980]: Relative Chebyshev centers in normed linear spaces I. J. Approximation Theory **29** (1980), 235-252

Bangert, V. [1981]: Totally convex sets in complete Riemannian manifolds. J. Diff. geom. **16** (1981), 333-345

Bonnesen, T., und Fenchel, W. [1934]: Theorie der konvexen Körper. I-VII u. 1-164. Springer Berlin **1934**. (Reprint Chelsea Publ. Comp. New York 1971)

Kobayashi, S. and Nomizu, K. [1963]: Foundations of differential geometry, Vol. I. i-xi a. 1-329. Interscience Publishers John Wiley and Sons **1963**

Phelps, R.R. [1957]: Convex sets and nearest points. Proc. Amer. Math. Soc. **8** (1957), 790-797

Wortmann, Ch. [2004]: Algorithmische Analyse von rohrförmigen Flächen für das Reverse Engineering. Thesis preprint Fachbereich Informatik Dortmund **2004**

Author's address:

Fakultät für Mathematik
Technische Universität Dortmund
D-44221 Dortmund, Germany
e-mail: rolf.walter@tu-dortmund.de